

Earth Syst. Sci. Data Discuss., referee comment RC1  
<https://doi.org/10.5194/essd-2020-348-RC1>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on **essd-2020-348**

Yang Song (Referee)

---

Referee comment on "Changes in global air pollutant emissions during the COVID-19 pandemic: a dataset for atmospheric modeling" by Thierno Doumbia et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2020-348-RC1>, 2021

---

The authors did an impressive work in collecting and analyzing activity data related to fossil fuel emissions during the COVID period. The results of this work can serve as a good reference in research related to the COVID impacts as well as atmospheric models. However, I believe that several issues could, and should, be further elaborated by the authors to help with the better understanding of this dataset.

1) line 109-112 says, "The collected data are then analyzed and an intercomparison of the changes in activity data from datasets providing similar or equivalent parameters are performed. The dataset that provides the most detailed and reliable data is then chosen."

This only applies to the transportation sector, correct? Because other sectors seem to suffer from lacking of data and the authors really didn't have much of a luxury to choose from different activity datasets for other sectors.

My question is, what is the standard for choosing the data in transportation sector? To be more specific, how did the authors define "reliable"?

The authors listed a comparison between the APPLE driving data and google mobility data. Figure S1 did show that APPLE data has more variations compared to google data. But does that mean APPLE data is less reliable? What is(are) the ultimate reason(s) for the discrepancies between APPLE data and google data? Besides the spatial coverage, what made the authors believe that google is more reliable than APPLE? Couldn't it be that in reality, the transportation emission IS that variant in different countries/regions?

2) Line 113-105, "The gridded daily/monthly files per sector are obtained by assigning the value of the AFs at the country/state/province level to each corresponding grid cell"

I have not found a detailed description on this in methodology. How exactly was the gridding implemented? The data sources listed in table 1 are mostly at national scale or province/state scale. How were they downscaled into 0.1 degree grids? What was used as the proxy? How reliable was the proxy?

3) Line 144-147, "In order to make the calculated AFs comparable with those derived using the other data sources considered in this study, the AFs for the Google's categories are scaled as a function of the Google mobility data, so that their values are less than 1 for a reduction in activity and above 1 otherwise."

I'm confused here. What is the baseline for 1? The authors mentioned in previous text in line 137-138 that "This baseline is calculated as the median value over the five-week period from January 3rd to February 6th 2020." But that was the google's definition of baseline for their mobility data, not the same baseline used to calculate AF in this paper, correct?

Or was it the same thing? Please clarify.

Also, is the baseline the same for all other sectors? Please clarify.

4) This is not a question but rather a comment. The "uncertainty" as showed in Figure 2,3,4,5,6 is not the result of validation based on other reliable datasets, but instead, it refers to the AF variations at different geographical locations. In my opinion, the authors should call it as what it is instead of insinuating that it could be something that it is not. I'm not saying the authors didn't do data validation, they did, as shown in figure S4 and S5. But the light pink color in figure 2,3,4,5,6 is not the result of validation.